Big data 101

İlk büyük veri projesi 1937'de Amerika'da başlamış.
Hükümet, 128 milyon kişinin sosyal güvenlik verilerini kaydetmiş. O tarihlerde bu iş, daha sonra IBM'e büyük ölçekli otomatik kayıt tutma ve modern veri analizi için temel oluşturacak endüstriyel kartlı makinelere dayanıyormuş.

Geleneksel şirketlerin yıllarca araştırma yaparak ortaya çıkarabildiği, çeşitli karmaşık bilgileri harmanlayıp bundan bir ürün oluşturma dönem artık bitti.
Eski zaman üretim hızı ve pratikleri ile böyle bir sürece girsek bile, ürün piyasaya çıktığında modası geçmiş olurdu.

Amazon, Facebook ve Google'ın veri kralları olduğu birinci sınıf dijital çağımızda güç, artık veriyi analiz etmeye dayalı. Bunun bir sonraki aşaması ise veriyi anlayarak, geleceğin trendini tahmin etmek ve bir adım öne geçmek. Tam da bu sebeplerle veri analizi altın madenciliğine benzetiliyor.

Günümüzde, web sitesi erişimleri ve tıklama ile ilgili her şey arşivlenir. Maliyetleri düşürmek ve rakiplerinden daha iyi performans elde etmek bu bilgiler zararsız bir şekilde kullanılır. Yine de büyük verinin totaliter kullanımları ile ilgili mutlaka bir kaç kara senaryo duymuşuzdur. Daha da ötesi, bu tür tehditlere dikkat çeken tonla yazı ve uzman görüşü olması olması.

Orwell'in 1984'ü gibi kurgusal distopyalar kadar olmasa da büyük verinin tehditkar kullanımları olasılıklar dahilinde.

Bazı görüşler, büyük verinin demokrasinin sonu olacağını söylüyor.
Cambridge Analytica şirketi, bireyleri anlamak için sosyal medya ağlarından gelen verileri inceleyerek hem AB referandumunu hem de daha sonra ABD seçimini etkileyebileceklerini gösterdiler.

Kimlerin büyük veriye sahip olduğu, kullandığı ve topladığı hakkında endişeler var. Toplanan bilgilerle ayrı olmayı vaat eden hükümetlere ve şirketlere rağmen, bilgilerin üçüncü şahıslar tarafından kullanılmayacağına dair sürekli güvence mevcut olsa da gerçekte çok farklı olabilir.

Geçtiğimiz aylarda, ISP ve internet şirketlerinin kişisel özel veriler satmasını engelleyen düzenlemeler ABD Senatosu tarafından kaldırıldı.

Bu durum, çevrimiçi gizliliği korumayı neredeyse imkânsız hale getiriyor. Aşırı kontrol sahibi hükümetlerin veya partilerin, genel nüfusu kendi bilgi ve servetlerine göre şekillendireceği olasılığı endişeleri arttırdı.

Çin'in pilot bölgelerde uygulamaya geçirdiği, sosyal kredi dereceleri ile vatandaşlarına ceza ve ödül sistemi belirlemesi endişeleri gerçek kılıyor. Jiangsu eyaletinde bir pilot program uygulandı: Standartın üstünde puan alanlar promosyonlarla ödüllendirildi ya da toplu konut sırasında öne geçtiler. Şimdilik, yaklaşık 30 yerel yönetim, bütün bir toplumun davranışını etkilemek için planlı bir şekilde veri topluyor.

Her gün, 2.5 Ekzabayt veri yaratıyoruz. (1milyon Terabyte, 1 Eksabayte yapıyor) Şu anda dünyadaki verilerin% 90'ı sadece son iki yılda yaratılmış. İklim bilgisi toplamak için kullanılan sensörler, sosyal medya sitelerine mesajlar, dijital resim ve videolar, satın alma işlem kayıtları ve cep telefonu GPS sinyalleri: Bu veriler büyük veridir.